Slowly Changing Dimensions Type 2:

In SCD Type2 incoming data record is identified as insert or update.This identification of record is done by primary key and surrogate key in the table. If the record is identified for update it is updated using flag, version number and effective date.

Mapping for SCD type2 using Update with Effective date.

In Target table two extra columns are created named Effective Start Date and Effective End Date. Also new column for surrogate key is added. Which has uniquely generated numbers by sequence generator transformation.

Step1:

Create source and target definitions by using appropriate DB connections.

Drag those definitions into mapping designer. Create three target instances to load records specified for insert and update.

Step2:

Write a SQL query override at source definition to select latest records from source.

Ex:Select EmoNo,EName,Sal from SRC_EMP where EED>(To_DATE('$$Set_Max_Date','MM/DD/YYYYHH24:MI:SS')

Step 3:

Create a Lookup Transformation and select Target table as base table for lookup to compare data with source table. Define the key columns (like Primary key and Surrogate key) to write lookup condition to compare source and target tables. Data retrieved based on the LKP condition will be cached in LKP Cache. Use LKP SQL override to retrieve active records from target.

Ex:Select Sur_Key,Emp_No,EName,Sal from TGT_EMP where Updated_date>SYSDATE

Step 4:

Create Expression Transformation, drag surrogate key and sal columns from lookup transformation and other columns from source qualifier transformation to Expression transformation.Create additional columns Effective start date ,effective end date and Inactive_Rec.Also set parameter (‘$$Set_Max_Date’) a value.Assign Inactive_Rec column as SYSDATE.

Also create a mapping variable for incremental load.new max value is to be assigned to mapping variable through Set_Max_Date variable.After sessions runs successfully new max value is set to this variable.

New ports and values:

Updated_Date: (‘$$Set_Max_Date’) to extract latest date.

Effective start date:SYSDATE

Effective end date: To_DATE(‘01/01/5000’,’MM/DD/YYYY’) to identify active and Inactive records.

Inactive_Rec:Sysdate to mark active status of record

Step 5:

Create Router transformation.Create conditions such that two groups are created one for insert and other for update with insert.

Group1(Insert):ISNULL(Surrogate_Key) new records are selected.

Group2(Insert with Update):NOT ISNULL(Surrogate_Key) and (Sal!=sal1) for updated record

Step 6:

Create Sequence generator transformation to generate unique values.Connect Next val port to Surrogate_Key port in target.

Step7:

Create Update strategy transformation to flag record for update .In Update strategy transformation select DD_UPDATE for records to be updated in target.

Step 8:

Connect ports from router insert group to target instance 1.

Connect Surrogate key port in target 3 to nextval in sequence generator.Ports from update strategy to target 3.

Connect required ports from Insert_Update group in router to target 2.

Below is final Mapping for SCD Type 2.